The Selection of High-Skilled Emigrants*
–Online and Data Appendix–

Matthias Parey, Jens Ruhose, Fabian Waldinger, and Nicolai Netz

*Matthias Parey: University of Essex and Institute for Fiscal Studies, m.parey@essex.ac.uk, Jens Ruhose: Leibniz University Hannover, ruhose@wipol.uni-hannover.de, Fabian Waldinger: London School of Economics, f.waldinger@lse.ac.uk, Nicolai Netz: DZHW, netz@dzhw.eu.

A Online Appendix

A.1 Appendix Tables and Figures

Table A.1: Empirical papers on the selection of international migrants

Paper	Skill measure	Home and destination countries	Consistent with basic Roy/Borjas model
Skill measure: actual earnings			
Borjas (1987)	Entry w in destination (U.S.)	Many countries -> U.S.	partly
Fernández-Huertas Moraga (2011)	w in home country (Mexico)	Mexico -> U.S.	yes
Kaestner and Malamud (2014)	w in home country (Mexico)	Mexico -> U.S.	yes
Borjas (2014)	Entry w in destination (U.S.)	Many countries -> U.S.	yes
Borjas, Kauppinen and Poutvaara (2015)	w in home country (Denmark)	Denmark -> many countries	yes
Skill measure: predicted earnings			
Ramos (1992)	w(edu, exp, marital status, occ)	Puerto Rico -> U.S.	yes
Chiquiar and Hanson (2005)	w(edu, age, marital status)	Mexico -> U.S.	no
Other skill measures:			
Chiquiar and Hanson (2005)	Education	Mexico -> U.S.	no
Orrenius and Zavodny (2005)	Education	Mexico -> U.S.	no
Feliciano (2005)	Education	32 countries -> U.S.	no
Ibarraran and Lubotsky (2007)	Predicted education	Mexico -> U.S.	yes
Borjas (2008)	Education	Puerto Rico -> U.S.	partly
Fernández-Huertas Moraga (2011)	Education	Mexico -> U.S.	yes
Grogger and Hanson (2011)	Education	Cross country	no
Abramitzky, Platt Boustan, Eriksson (2012)	Father's occupation	Norway -> U.S.	yes
	Own occupation		partly
Stolz and Baten (2012)	Age heaping	52 countries -> 5 countries	yes
Belot and Hatton (2012)	Education	80 countries -> 29 countries	partly
Kaestner and Malamud (2014)	Education	Mexico -> U.S.	no
	Cognitive ability		no
Borjas, Kauppinen and Poutvaara (2015)	Residual wages	Denmark -> many countries	yes
Gould and Moav (2016)	Education	Israel-> U.S.	yes
	Residual wages		partly

Notes: The table shows empirical tests of the Roy/Borjas model ordered by skill measure and year of publication. Papers that report selection for different skill measures are listed more than once.

Table A.2: Destinations of German university graduates

Country	Number of graduates	Wage inequality data
Germany	10,510	Yes
Switzerland	152	Yes
United States	87	Yes
UK	68	Yes
Austria	42	Yes
France	41	Yes
Luxembourg	25	Yes
Netherlands	25	Yes
Spain	20	Yes
Belgium	20	Yes
Norway	20	Yes
Sweden	15	Yes
Italy	13	Yes
Denmark	13	Yes
Ireland	11	Yes
China	8	No
Australia	7	Yes
Canada	7	Yes
Japan	5	Yes
Finland	5	Yes
Poland	5	Yes
Brazil	5	No
New Zealand	5	No
Other	56	No

Notes: The table shows the most important destinations of German university graduates in the graduate survey data and the availability of inequality data for university graduates in the augmented LIS data. All destinations in the category "Other" receive less than five graduates.

Table A.3: Data sources on earnings inequality by country

Country	***Data Source***	***Earnings***	***Currency***	***Years (Observations)***
Australia	LIS	gross	Australian Dollar	1995 (*1,010*); 2001 (*752*); 2003 (*1,170*)
Austria	LIS	net	Schilling	1997 (*121*); 2000 (*97*)
	LIS	gross	Euro	2004 (*372*)
	Microcensus	net	Schilling	1999 (*1,394*)
	EU-SILC	net and gross	Euro	2007 (*597*); 2008 (*654*)
Belgium	LIS	gross	Belgian franc	1997 (*215*)
		net		2000 (*520*)
Canada	LIS	gross	Canadian Dollar	1998 (*3,236*); 2000 (*2,908*); 2004 (*3,324*); 2007 (*3,669*); 2010 (*3,791*)
Denmark	LIS	gross	Danish Krone	1995 (*9,451*); 2000 (*13,464*); 2004 (*15,307*)
Finland	LIS	gross	Finnish Markka	1995 (*1,275*); 2000 (*1,726*)
			Euro	2004 (*6,633*); 2007 (*1,793*); 2010 (*1,683*)
France	LIS	net	French Franc	1994 (*728*); 2000 (*675*)
			Euro	2005 (*1,115*)
Germany	LIS	gross	Deutsche Mark	1994 (*822*); 2000 (*1,511*)
			Euro	2004 (*1,622*); 2007 (*1,529*); 2010 (*1,536*)
Ireland	LIS	net	Irish Pound	1996 (*150*); 2000 (*148*)
		gross	Euro	2004 (*610*); 2007 (*748*); 2010 (*505*)
Italy	LIS	net	Italian Lira	1998 (*418*); 2000 (*493*)
			Euro	2004 (*440*); 2008 (*633*); 2010 (*739*)
Japan	LIS	gross	Yen	2008 (*841*)
Luxembourg	LIS	net	Luxembourg Franc	1997 (*297*); 2000 (*391*)
		gross	Euro	2004 (*720*); 2007 (*780*); 2010 (*953*)
Netherlands	LIS	gross	Netherlands Guilder	1993 (*687*); 1999 (*772*)
			Euro	2004 (*1,801*); 2007 (*2,127*); 2010 (*2,236*)

(continued on next page)

Table A.3 (continued)

Country	***Data Source***	***Earnings***	***Currency***	***Years (Observations)***
Norway	LIS	gross	Norwegian Krone	1995 (*2,031*); 2000 (*3,501*); 2004 (*3,920*)
Poland	LIS	net	Zloty	1995 (*2,642*)
		gross		2007 (*4,381*); 2010 (*6,358*)
Spain	LIS	net	Spanish Peseta	1995 (*579*); 2000 (*528*)
		net	Euro	2004 (*2,318*)
		gross	Euro	2007 (*2,175*); 2010 (*2,138*)
Sweden	LIS	gross	Swedish Krona	1995 (*2,427*); 2000 (*3,115*); 2005 (*2,605*)
Switzerland	SAKE	net and gross	Swiss Franc	1998-2005 (*2,394*)
UK	LIS	gross	Pound Sterling	1995 (*459*); 1999 (*2,840*); 2004 (*3,379*); 2007 (*3,233*); 2010 (*3,610*)
United States	LIS	gross	US Dollar	1997 (*12,988*); 2000 (*13,443*); 2004 (*23,229*); 2007 (*24,295*); 2010 (*24,026*)

Notes: Unweighted number of observations of university graduates, 30 to 60 years old, working in full-time dependent employment, with strictly positive earnings are reported in parentheses. For Switzerland, we report the average number of observations over all years. The main data source is the Luxembourg Income Study (LIS). For Austria and Switzerland, we use additional surveys. Austria: Microcensus (1999) and Survey on Income and Living Conditions (EU-SILC, 2007, 2008). Switzerland: Swiss Labour Force Survey (SAKE, 1998-2005). LIS contains different surveys for the countries: Australia: Survey of Income and Housing Costs (SIHC); Austria: European Household Panel (ECHP, 1997, 2000), EU-SILC (2004); Belgium: Socio-Economic Panel (SEP, 1997), Panel Study of Belgian Households (PSBH, 2000); Canada: Survey of Labour and Income Dynamics (SLID); Denmark: Income Tax Register; Finland: Income Distribution Survey (IDS, 1995, 2000, 2004), EU-SILC (2007, 2010); France: Family Budget Survey (BdF); Germany: German Socio-Economic Panel (GSOEP); Ireland: Living in Ireland Survey (ECHP, 1996, 2000), EU-SILC (2004, 2007, 2010); Italy: Survey on Household Income and Wealth (SHIW); Japan: Japan Household Panel Survey (JHPS); Luxembourg: Socio-Economic Panel "Living in Luxembourg" (PSELL); Netherlands: Socio-Economic Panel Survey (1999, 2004), EU-SILC (2004, 2007, 2010); Norway: Income Distribution Survey (IF); Poland: Household Budget Survey; Spain: Spanish European Community Household Panel (ECHP, 1995, 2000), EU-SILC (2004, 2007, 2010); Sweden: Income Distribution Survey (HINK); UK: Family Resources Survey (FRS); United States: Current Population Survey (CPS).

Table A.4: Inequality measures

	(1)	(2)	(3)	(4)	(5)	(6)
Country	75/25 ratio graduates	90/50 ratio overall population	75/25 ratio overall population	90/10 ratio, OECD overall population	Gini, OECD overall population	Theil index overall population
United States	1.930	2.085	2.070	5.840	0.362	0.161
France	1.889	1.890	1.736	3.480	0.288	0.097
Poland	1.873	1.960	1.884	4.414	0.322	0.121
Italy	1.806	1.639	1.445	4.220	0.317	0.146
Spain	1.766	1.877	1.790	4.657	0.316	0.130
Japan	1.749	1.742	2.084	5.025	0.326	0.115
Canada	1.733	1.872	1.954	4.062	0.323	0.123
United Kingdom	1.724	1.961	1.866	4.233	0.340	0.134
Austria	1.650	1.717	1.721	3.400	0.263	0.099
Luxembourg	1.553	1.788	1.915	3.388	0.275	0.102
Switzerland	1.551	1.663	1.598	3.600	0.290	0.100
Belgium	1.540	1.580	1.467	3.329	0.261	0.079
Germany	1.524	1.625	1.476	3.460	0.276	0.097
Ireland	1.521	1.686	1.596	3.886	0.307	0.100
Sweden	1.467	1.497	1.427	3.080	0.255	0.069
Netherlands	1.450	1.590	1.406	3.300	0.290	0.082
Australia	1.439	1.622	1.501	4.350	0.310	0.099
Norway	1.409	1.514	1.461	2.880	0.262	0.070
Finland	1.395	1.581	1.342	3.062	0.256	0.076
Denmark	1.347	1.437	1.314	2.786	0.230	0.056
Correlation with 75/25 ratio (graduates)	1.000	0.864	0.755	0.736	0.748	0.839

Notes: The inequality measures reported in column (1) (75/25 ratio graduates) are computed from a sample of university graduates, working full-time, 30-60 years old, based on net earnings. Inequality measures reported in columns (2) to (6) are computed for the overall population. All inequality measures are averaged for the time period 1998-2010. Data for the measures reported in columns (1) to (3) and (6) come from the Luxembourg Income Study (LIS) for most countries, from the Microcensus and EU-SILC for Austria, and from SAKE for Switzerland. Data on the 90/10 ratio and Gini coefficients reported in column (4) and (5) come from the OECD. See Data Appendix **??** for details.

Table A.5: Stochastic dominance tests of reverse hypotheses

	Test statistic	p-value: Kolmogorov-Smirnov	p-value: Barrett-Donald
	(1)	(2)	(3)
Panel A: Selection to *more equal* and to *less equal* destinations			
Panel A1: OLS			
'Home' vs 'Equal'	-0.014	0.963	0.988
'Unequal' vs 'Home'	0.000	1.000	1.000
'Unequal' vs 'Equal'	-0.004	0.997	0.998
Panel A2: Heckman selection correction			
'Home' vs 'Equal'	-0.015	0.956	0.971
'Unequal' vs 'Home'	0.000	1.000	1.000
'Unequal' vs 'Equal'	-0.004	0.997	1.000
Panel B: Selection to *very equal*, to *somewhat equal*, to *somewhat unequal*, and to *very unequal* destinations			
Panel B1: OLS			
Stochastic dominance tests for very equal and very unequal destinations			
'Home' vs 'Very equal'	-0.016	0.981	0.990
'Very unequal' vs 'Home'	-0.010	0.975	0.993
'Very unequal' vs 'Very equal'	0.000	1.000	1.000
Stochastic dominance tests for more similar destinations			
'Somewhat equal' vs 'Very equal'	-0.018	0.985	0.992
'Home' vs 'Somewhat equal'	-0.049	0.758	0.882
'Somewhat unequal' vs 'Home'	-0.008	0.952	0.992
'Very unequal' vs 'Somewhat unequal'	-0.030	0.843	0.949
Panel A2: Heckman selection correction			
Stochastic dominance tests for very equal and very unequal destinations			
'Home' vs 'Very equal'	-0.018	0.977	0.976
'Very unequal' vs 'Home'	-0.010	0.975	0.991
'Very unequal' vs 'Very equal'	0.000	1.000	0.999
Stochastic dominance tests for more similar destinations			
'Somewhat equal' vs 'Very equal'	-0.009	0.996	0.999
'Home' vs 'Somewhat equal'	-0.052	0.736	0.829
'Somewhat unequal' vs 'Home'	-0.005	0.981	0.988
'Very unequal' vs 'Somewhat unequal'	-0.030	0.845	0.964

Notes: The table reports one-sided Kolmogorov-Smirnov test statistics and Kolomogorov-Smirnov and Barrett and Donald p-values for CDFs reported in Figure 2. See Table 3 for details.

Table A.6: Stochastic dominance tests for sample with imputations

	Test statistic	p-value: Kolmogorov-Smirnov		p-value: Barrett-Donald	
	(1)	(2)		(3)	
Panel A: Selection to *more equal* and to *less equal* destinations					
Panel A1: OLS					
'Equal' vs 'Home'	0.145	0.004	***	0.017	**
'Home' vs 'Unequal'	0.052	0.036	**	0.109	
'Equal' vs 'Unequal'	0.173	0.001	***	0.005	***
Panel A2: Heckman selection correction					
'Equal' vs 'Home'	0.137	0.007	***	0.049	**
'Home' vs 'Unequal'	0.059	0.015	**	0.099	*
'Equal' vs 'Unequal'	0.172	0.001	***	0.011	***
Panel B: Selection to *very equal*, to *somewhat equal*, to *somewhat unequal*, and to *very unequal* destinations					
Panel B1: OLS					
Stochastic dominance tests for very equal and very unequal destinations					
'Very equal' vs 'Home'	0.176	0.052	*	0.097	*
'Home' vs 'Very unequal'	0.122	0.007	***	0.035	**
'Very equal' vs 'Very unequal'	0.241	0.013	**	0.030	**
Stochastic dominance tests for more similar destinations					
'Very equal' vs 'Somewhat equal'	0.116	0.440		0.621	
'Somewhat equal' vs 'Home'	0.128	0.061	*	0.142	
'Home' vs 'Somewhat unequal'	0.048	0.123		0.267	
'Somewhat unequal' vs 'Very unequal'	0.110	0.050	*	0.134	
Panel B2: Heckman selection correction					
Stochastic dominance tests for very equal and very unequal destinations					
'Very equal' vs 'Home'	0.172	0.060	*	0.154	
'Home' vs 'Very unequal'	0.134	0.003	***	0.032	**
'Very equal' vs 'Very unequal'	0.247	0.011	**	0.029	**
Stochastic dominance tests for more similar destinations					
'Very equal' vs 'Somewhat equal'	0.116	0.440		0.638	
'Somewhat equal' vs 'Home'	0.124	0.075	*	0.217	
'Home' vs 'Somewhat unequal'	0.056	0.056	*	0.200	
'Somewhat unequal' vs 'Very unequal'	0.116	0.037	**	0.134	

Notes: The table reports one-sided Kolmogorov-Smirnov test statistics and Kolmogorov-Smirnov and Barrett and Donald p-values for CDFs reported in Online Appendix Figure A.3. Online Appendix **??** reports details on the imputation procedure. See Table 3 for details. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.7: Stochastic dominance tests for quantile regression analysis

	Test statistic	p-value Barrett-Donald	
	(1)	(2)	
Panel A: Controlling for log net earnings			
Stochastic dominance tests for very equal and very unequal destinations			
'Very equal' vs 'Home'	0.232	0.076	*
'Home' vs 'Very unequal'	0.131	0.045	**
'Very equal' vs 'Very unequal'	0.283	0.033	**
Stochastic dominance tests for more similar destinations			
'Very equal' vs 'Somewhat equal'	0.192	0.384	
'Somewhat equal' vs 'Home'	0.152	0.280	
'Home' vs 'Somewhat unequal'	0.030	0.912	
'Somewhat unequal' vs 'Very unequal'	0.192	0.024	**
Panel B: Controlling for family expenditure			
Stochastic dominance tests for very equal and very unequal destinations			
'Very equal' vs 'Home'	0.222	0.132	
'Home' vs 'Very unequal'	0.141	0.039	**
'Very equal' vs 'Very unequal'	0.283	0.034	**
Stochastic dominance tests for more similar destinations			
'Very equal' vs 'Somewhat equal'	0.182	0.445	
'Somewhat equal' vs 'Home'	0.141	0.300	
'Home' vs 'Somewhat unequal'	0.071	0.143	
'Somewhat unequal' vs 'Very unequal'	0.131	0.117	
Panel C: Controlling for the unemployment rate			
Stochastic dominance tests for very equal and very unequal destinations			
'Very equal' vs 'Home'	0.283	0.010	**
'Home' vs 'Very unequal'	0.141	0.041	**
'Very equal' vs 'Very unequal'	0.303	0.014	**
Stochastic dominance tests for more similar destinations			
'Very equal' vs 'Somewhat equal'	0.172	0.409	
'Somewhat equal' vs 'Home'	0.162	0.196	
'Home' vs 'Somewhat unequal'	0.040	0.690	
'Somewhat unequal' vs 'Very unequal'	0.141	0.096	*

(continued on next page)

Table A.7 (continued)

	Test statistic	p-value: Barrett-Donald	
	(1)	(2)	
Panel D: Controlling for life satisfaction			
Stochastic dominance tests for very equal and very unequal destinations			
'Very equal' vs 'Home'	0.303	0.005	***
'Home' vs 'Very unequal'	0.172	0.010	**
'Very equal' vs 'Very unequal'	0.354	0.001	***
Stochastic dominance tests for more similar destinations			
'Very equal' vs 'Somewhat equal'	0.192	0.320	
'Somewhat equal' vs 'Home'	0.141	0.268	
'Home' vs 'Somewhat unequal'	0.051	0.446	
'Somewhat unequal' vs 'Very unequal'	0.162	0.045	**
Panel E: All controls			
Stochastic dominance tests for very equal and very unequal destinations			
'Very equal' vs 'Home'	0.202	0.228	
'Home' vs 'Very unequal'	0.131	0.126	
'Very equal' vs 'Very unequal'	0.253	0.160	
Stochastic dominance tests for more similar destinations			
'Very equal' vs 'Somewhat equal'	0.202	0.360	
'Somewhat equal' vs 'Home'	0.152	0.307	
'Home' vs 'Somewhat unequal'	0.030	0.922	
'Somewhat unequal' vs 'Very unequal'	0.182	0.078	*

Notes: The table reports one-sided Kolmogorov-Smirnov test statistics and Barrett and Donald p-values for CDFs reported in Figure A.7. See Table 3 for details. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: Stochastic dominance tests for alternative inequality measures

		p-value			
	Test statistic	Kolmogorov-Smirnov		Barrett-Donald	
	(1)	(2)		(3)	
Panel A: OLS					
Panel A1: 75/25 ratio, university graduates (baseline)					
'Equal' vs 'Home'	0.187	0.001	***	0.006	***
'Home' vs 'Unequal'	0.061	0.031	**	0.098	*
'Equal' vs 'Unequal'	0.220	0.000	***	0.001	***
Panel A2: 90/50 ratio, overall population					
'Equal' vs 'Home'	0.204	0.000	***	0.001	***
'Home' vs 'Unequal'	0.066	0.019	**	0.068	*
'Equal' vs 'Unequal'	0.247	0.000	***	0.000	***
Panel A3: 75/25 ratio, overall population					
'Equal' vs 'Home'	0.191	0.000	***	0.001	***
'Home' vs 'Unequal'	0.071	0.011	**	0.049	**
'Equal' vs 'Unequal'	0.235	0.000	***	0.000	***
Panel A4: 90/10 ratio, overall population					
'Equal' vs 'Home'	0.101	0.037	**	0.102	
'Home' vs 'Unequal'	0.066	0.031	**	0.092	*
'Equal' vs 'Unequal'	0.138	0.011	**	0.030	**
Panel A5: Gini coefficient, overall population					
'Equal' vs 'Home'	0.095	0.085	*	0.196	
'Home' vs 'Unequal'	0.059	0.053	*	0.135	
'Equal' vs 'Unequal'	0.124	0.038	**	0.101	
Panel A6: Theil index, overall population					
'Equal' vs 'Home'	0.183	0.000	***	0.001	***
'Home' vs 'Unequal'	0.072	0.012	**	0.047	**
'Equal' vs 'Unequal'	0.235	0.000	***	0.000	***

(continued on next page)

Table A.8 (continued)

	Test statistic	p-value: Kolmogorov-Smirnov		p-value: Barrett-Donald	
	(1)	(2)		(3)	
Panel B: Heckman selection correction					
Panel B1: 75/25 ratio, university graduates (baseline)					
'Equal' vs 'Home'	0.182	0.002	***	0.022	**
'Home' vs 'Unequal'	0.071	0.009	***	0.083	*
'Equal' vs 'Unequal'	0.218	0.000	***	0.004	***
Panel B2: 90/50 ratio, overall population					
'Equal' vs 'Home'	0.197	0.000	***	0.015	**
'Home' vs 'Unequal'	0.075	0.005	***	0.056	*
'Equal' vs 'Unequal'	0.240	0.000	***	0.001	***
Panel B3: 75/25 ratio, overall population					
'Equal' vs 'Home'	0.188	0.000	***	0.009	***
'Home' vs 'Unequal'	0.081	0.003	***	0.045	**
'Equal' vs 'Unequal'	0.233	0.000	***	0.001	***
Panel B4: 90/10 ratio, overall population					
'Equal' vs 'Home'	0.096	0.052	*	0.233	
'Home' vs 'Unequal'	0.072	0.016	**	0.088	*
'Equal' vs 'Unequal'	0.135	0.014	**	0.053	*
Panel B5: Gini coefficient, overall population					
'Equal' vs 'Home'	0.094	0.087	*	0.292	
'Home' vs 'Unequal'	0.067	0.021	**	0.099	*
'Equal' vs 'Unequal'	0.127	0.033	**	0.090	*
Panel B6: Theil index, overall population					
'Equal' vs 'Home'	0.178	0.000	***	0.004	***
'Home' vs 'Unequal'	0.083	0.003	***	0.049	**
'Equal' vs 'Unequal'	0.232	0.000	***	0.000	***

Notes: The table reports one-sided Kolmogorov-Smirnov test statistics and Kolomogorov-Smirnov and Barrett and Donald p-values for CDFs reported in Online Appendix Figure A.8. See Table 3 for details. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.9: Stochastic dominance tests for selected destinations

		p-value			
	Test statistic	Kolmogorov-Smirnov		Barrett-Donald	
	(1)	(2)		(3)	
Panel A: Selection to *EU countries (2005), Norway,* and *Switzerland*					
Panel A1: OLS					
'Equal' vs 'Home'	0.206	0.001	***	0.003	***
'Home' vs 'Unequal'	0.052	0.133		0.289	
'Equal' vs 'Unequal'	0.231	0.000	***	0.002	***
Panel A2: Heckman selection correction					
'Equal' vs 'Home'	0.201	0.001	***	0.016	**
'Home' vs 'Unequal'	0.061	0.061	*	0.187	
'Equal' vs 'Unequal'	0.231	0.000	***	0.003	***
Panel B: Selection to *Austria* and *Switzerland*					
Panel B1: OLS					
'Home' vs 'Austria/Switzerland'	0.095	0.031	**	0.090	*
Panel B2: Heckman selection correction					
'Home' vs 'Austria/Switzerland'	0.102	0.019	**	0.075	*
Panel C: Selection to the *United States*					
Panel C1: OLS					
'Home' vs 'United States'	0.173	0.006	***	0.033	**
Panel C2: Heckman selection correction					
'Home' vs 'United States'	0.191	0.002	***	0.022	**

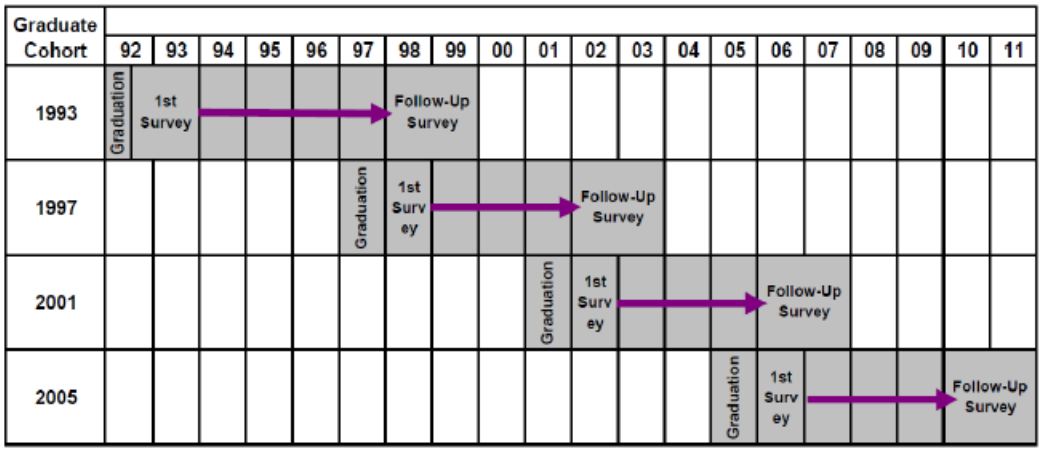
Notes: The table reports one-sided Kolmogorov-Smirnov test statistics and Kolomogorov-Smirnov and Barrett and Donald p-values for CDFs presented in Figures 4 and 6(a). See Table 3 for details. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.10: Mincer regressions for U.S. native graduates and German graduates in the United States

	(1)	(2)
Dependent variable: Log labor earnings	US Natives	Germans
University degree		
Doctoral degree	0.286***	0.305***
	(0.006)	(0.061)
Professional degree beyond a bachelor's degree	0.550***	0.563***
	(0.005)	(0.096)
Master's degree	0.184***	0.357***
	(0.002)	(0.047)
Age *(relative to sample mean)*		
Age	0.020***	0.039**
	(0.0009)	(0.018)
Age squared	–0.001***	–0.004**
	(0.0001)	(0.002)
Female		
Female	–0.224***	–0.349***
	(0.002)	(0.045)
Partner/Children		
Married	0.109***	0.059
	(0.003)	(0.050)
Children	0.044***	0.038
	(0.002)	(0.049)
Constant	10.889***	10.936***
	(0.011)	(0.186)
Graduate cohort FE	YES	YES
Subject FE	YES	YES
R-sq.	0.222	0.278
Observations	289,538	985

Notes: The table shows Mincer regressions for U.S. natives in column (1) and for Germans (in the United States) in column (2). For this regression we use Germans who were born in Germany to non-U.S. parents and moved to the United States at any point in their life. See Data Appendix **??** for details on the sample construction and data source. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

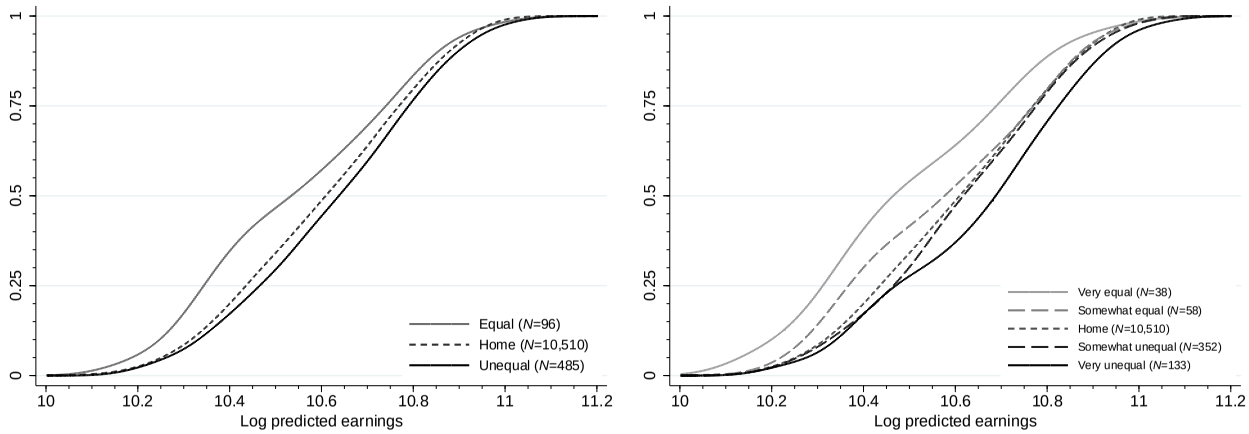
Figure A.1: Design of DZHW graduate panels



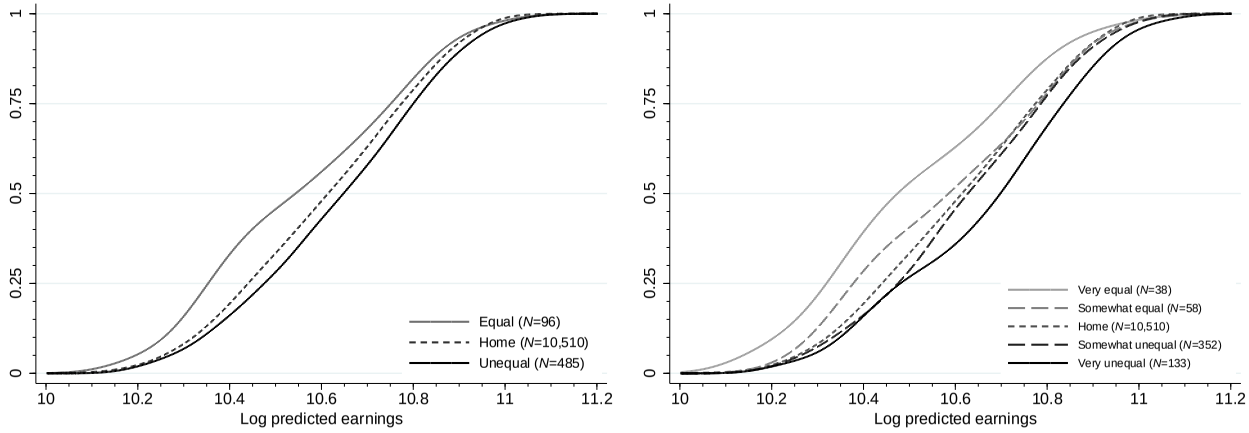
Notes: The figure shows the timing of the baseline and the five-year follow-up surveys of the DZHW Graduate Panels.

Figure A.2: Predicted earnings of migrants and non-migrants – Smoothed

(a) CDF for three and five groups of countries



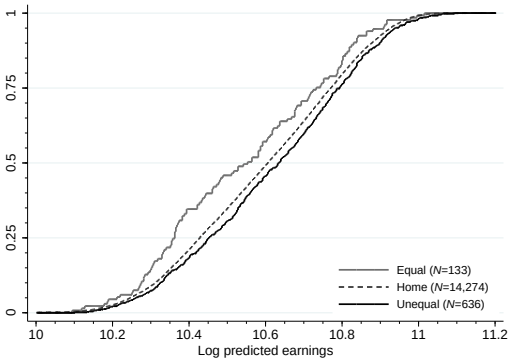
(b) CDF for three and five groups of countries – Earnings prediction corrected for selection



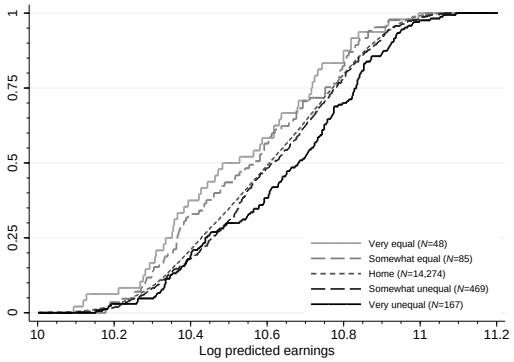
Notes: Figure shows kernel smoothed versions of the CDFs in Figure 2. Smoothed CDFs are based on the Gaussian kernel. We choose the bandwidth separately for each migrant group to account for the differences in the corresponding sample sizes. Bandwidth is chosen according to Silverman's rule of thumb (Silverman, 1986, p. 48), which we then rescale with a factor of 0.6 to avoid over-smoothing.

Figure A.3: Sensitivity: Addressing attrition

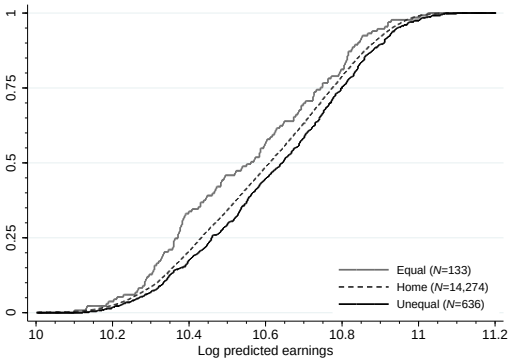
(a) Three groups – OLS



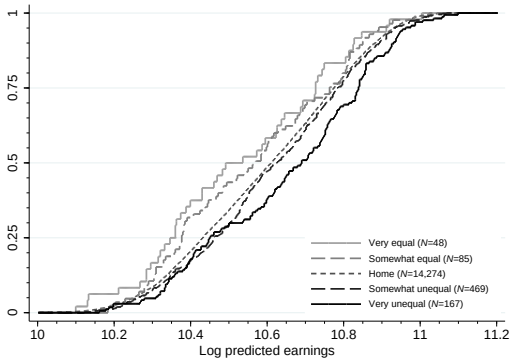
(b) Five groups – OLS



(c) Three groups – Heckman correction



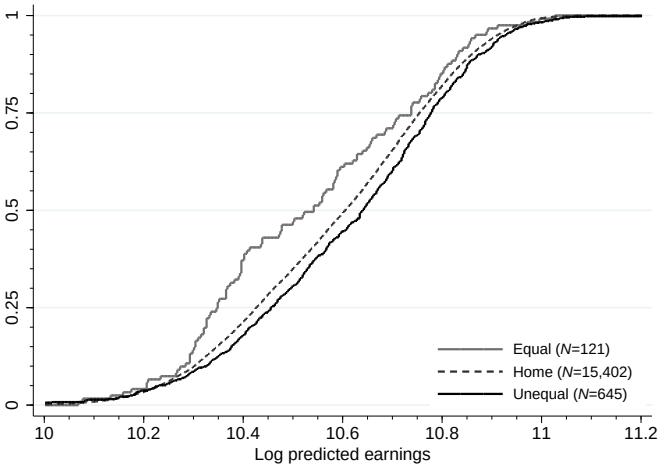
(d) Five groups – Heckman correction



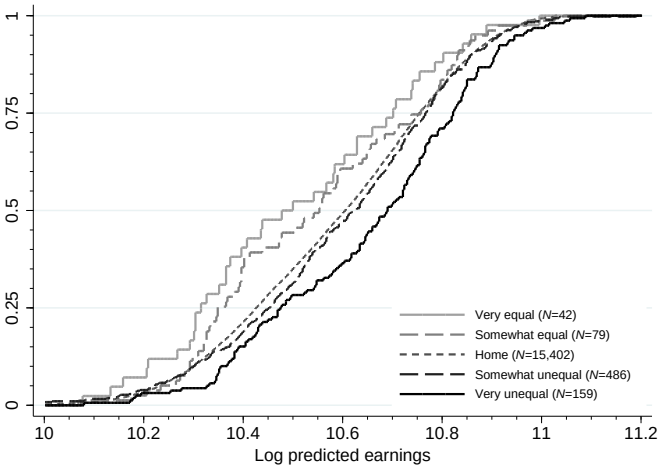
Notes: The figure shows CDFs of predicted earnings for three groups of countries (panels (a) and (c)) and for five groups of countries (panels (b) and (d)). The earnings prediction is based on a Mincer earnings regression on the full sample of graduates who respond in the initial survey. In panels (c) and (d), we use a prediction based on selection-corrected returns. Online Appendix **??** reports details on the imputation procedure and Online Appendix Table A.6 reports stochastic dominance tests. Corresponding baseline results are found in Figure 2.

Figure A.4: Sensitivity: Including graduates from non-traditional universities

(a) CDF for three groups of countries



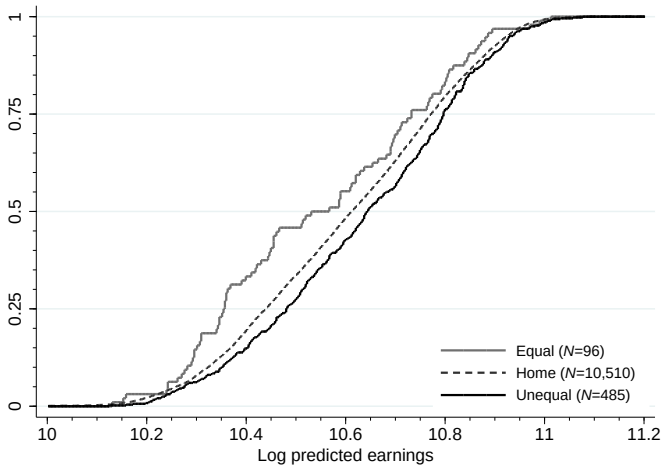
(b) CDF for five groups of countries



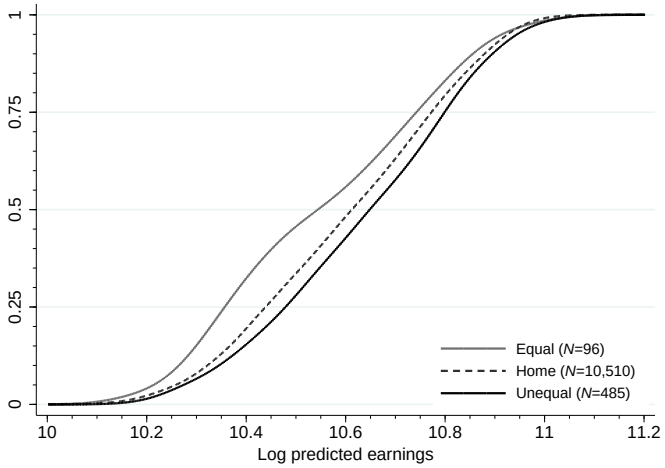
Notes: The figure shows CDFs of predicted earnings for migrants and non-migrants. In addition to graduates from traditional universities, the sample also consists of graduates from universities of applied sciences (*Fachhochschulen*), specialized universities focusing on arts, music, or theology, and private universities. Corresponding baseline results are found in Figure 2.

Figure A.5: Sensitivity: Mincer regression excluding presence of children and partnership status

(a) CDF



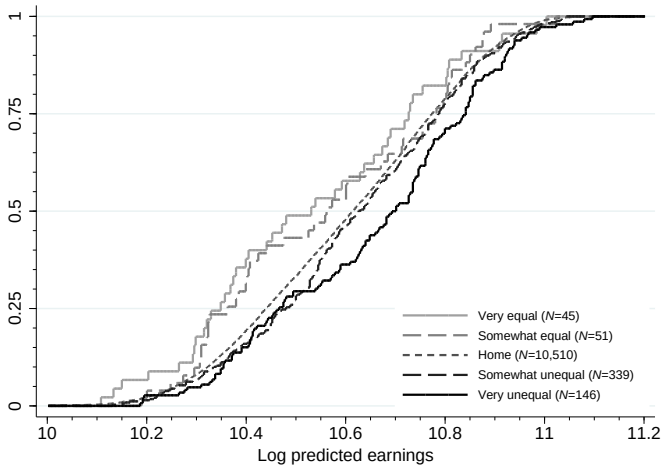
(b) CDF - smoothed



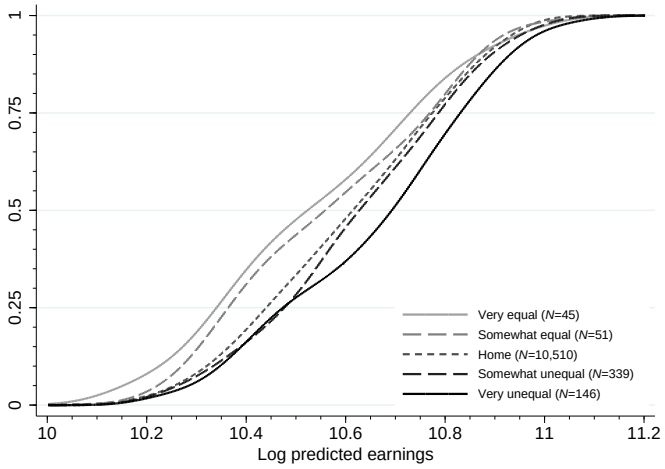
Notes: The figure shows a sensitivity analysis for the CDF results when we drop presence of children and marital/partnership status covariates from the Mincer regression. The figures correspond to the case with sample selection correction, corresponding baseline results are found in Figure 2.

Figure A.6: Sensitivity: Alternative definition of most equal and most unequal countries

(a) CDF

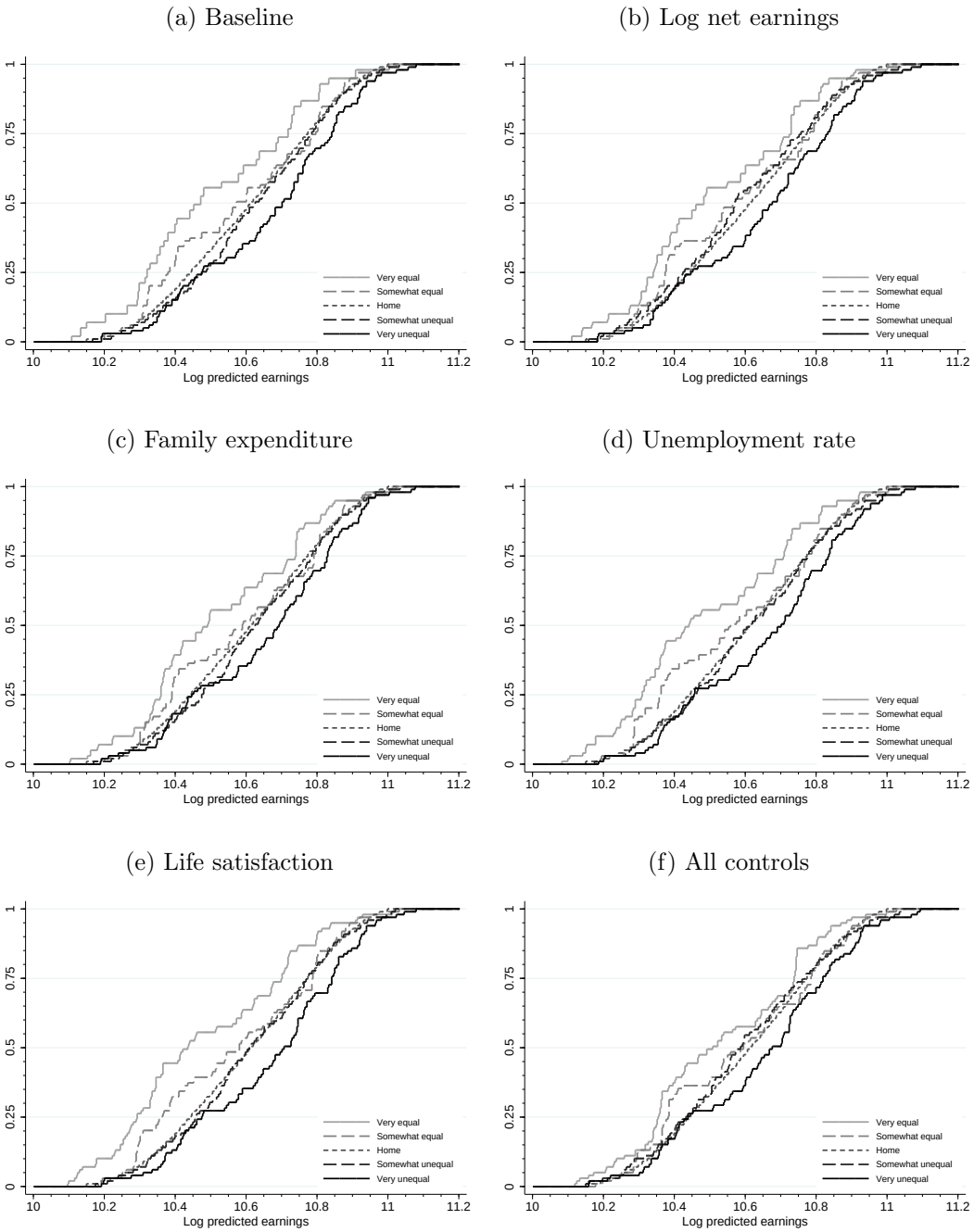


(b) CDF - smoothed



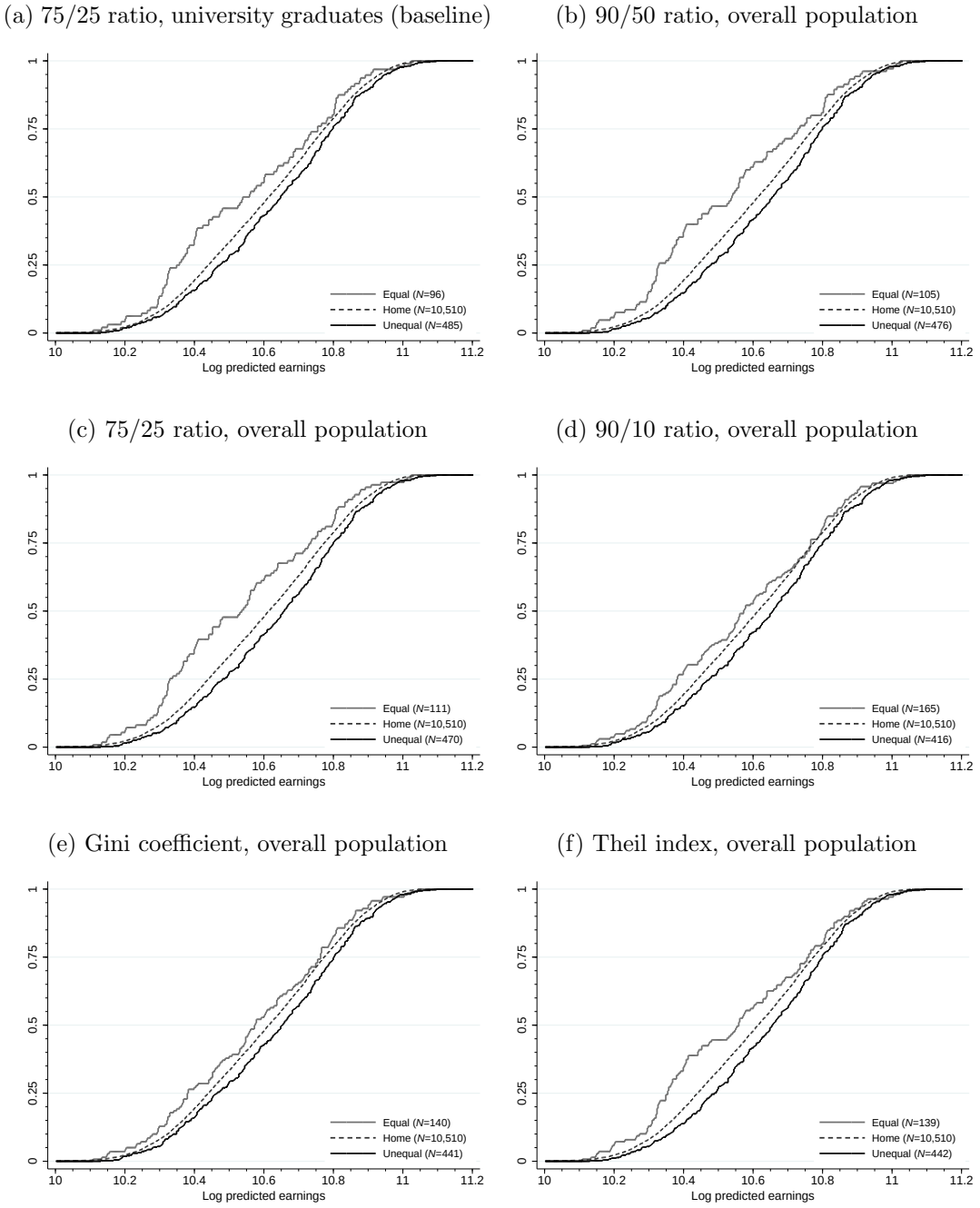
Notes: The figure shows a sensitivity analysis for the CDF results when we define most equal (most unequal) as the four countries with the lowest (highest) inequality index. The figures correspond to the case with sample selection correction. Corresponding baseline results are found in Figure 2.

Figure A.7: CDFs adjusted for potential confounders



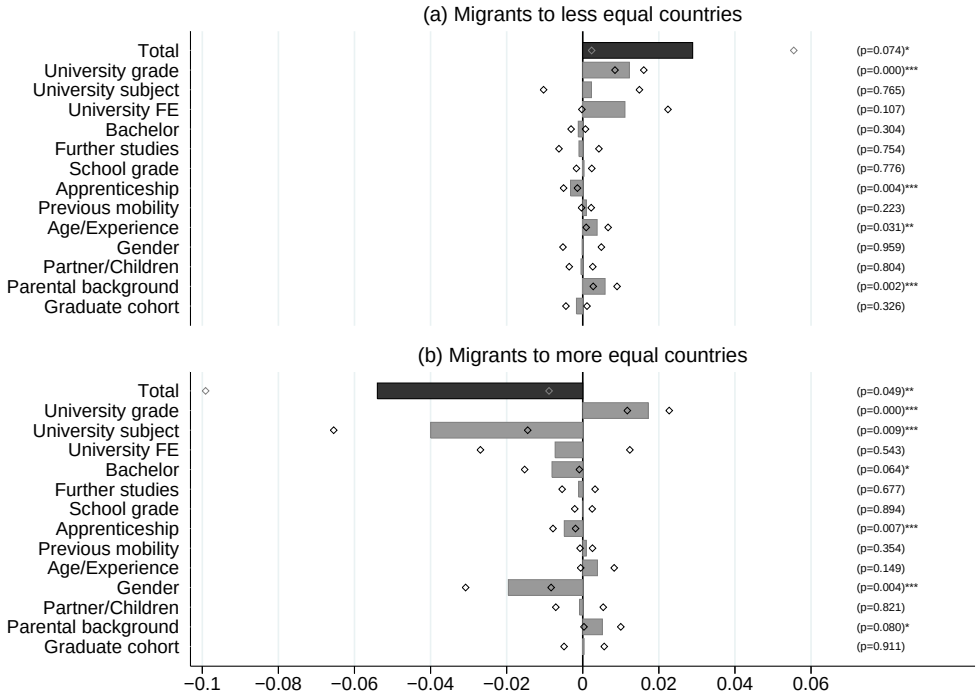
Notes: The figure shows adjusted CDFs of predicted earnings (from the selection-corrected Mincer regression) based on quantile regressions. Panel (a) replicates the baseline CDFs from Figure 2(b). In panel (b), we control for log net earnings, in panel (c) for public expenditures on family benefits as percentage of GDP, in panel (d) for the tertiary-educated unemployment rate, in panel (e) for life satisfaction, and in panel (f) for all controls. See Data Appendix **??** for details on data sources and Data Appendix Table B.2 for country-level controls. Online Appendix Table A.7 reports stochastic dominance tests. Results that are based on predicted earnings from the uncorrected Mincer regression are very similar. The derived CDFs from the quantile regressions occasionally violate local monotonicity. We therefore apply the method from Chernozhukov et al. (2010). This procedure ensures monotonicity in the CDFs; the estimates without this procedure are very similar.

Figure A.8: Sensitivity: Alternative inequality measures



Notes: Panel (a) shows CDFs of predicted earnings (prediction based on selection-corrected returns reported in column (2) of Table 2) for three groups: migrants to more equal countries, non-migrants, and migrants to less equal countries. Countries are classified as either more or less equal using the 75/25 ratio for university graduates (baseline). In panel (b), countries are classified using the 90/50 ratio for the overall population. In panel (c), countries are classified using the 75/25 ratio for the overall population. In panel (d), countries are classified using the 90/10 ratio for the overall population reported by the OECD. In panel (e), countries are classified using the Gini coefficient for the overall population. In panel (f), countries are classified using the Theil index for the overall population (see Data Appendix **??** for the construction of the Theil index). Online Appendix Table A.4 reports inequality measures. Online Appendix Table A.8 reports stochastic dominance tests.

Figure A.9: Decomposition of predicted earnings using selection-corrected coefficients



Notes: Panel (a) decomposes the mean difference in predicted earnings between migrants to less equal countries and non-migrants using selection-corrected Mincer regression coefficients (Table 2, column (2)). The top bar (black) measures the total difference in predicted earnings. The other bars decompose the total difference into the contributions of groups of characteristics (e.g. university grade). See Figure 5 for details. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

A.2 Bootstrap Test for Differences between the CDFs

In this section, we describe the bootstrap procedure that we implement to test differences between CDFs of earnings potential. We adapt the method developed by Barrett and Donald (2003) to our application. To facilitate the exposition, we denote the number of stayers by N^{Home} , the number of individuals observed in less equal destinations by N^{Unequal} , and the number of individuals observed in more equal destinations by N^{Equal}. We use the following bootstrap procedure:

1. Draw a sample of size N^{Home} from the sample of stayers, with replacement. Similarly, draw samples of sizes N^{Unequal} and N^{Equal} from the migrants observed in less and in more equal destinations, respectively. These data form the bootstrap sample b, which we denote with a star (*).

2. Use the bootstrap sample to estimate the Mincer wage regression, resulting in a coefficient estimate of $\hat{\beta}^\star$. Predict earnings potential for every observation in the bootstrap sample, $\hat{\theta}^\star_{0i}$. Construct the corresponding CDFs of earnings potential in each destination,

$$F^\star(\hat{\theta}^\star_0 \,|\, \text{Migration status}). \tag{A.1}$$

3. As an example, we focus on the test between 'Home' and 'Unequal' destinations. Following equation (11) in Barrett and Donald (2003), construct the test statistic

$$\bar{S}_b^{\text{Unequal, Home}} = \sup_z \left[(F^\star(z \,|\, \text{Home}) - F^\star(z \,|\, \text{Unequal})) - (F(z \,|\, \text{Home}) - F(z \,|\, \text{Unequal})) \right]. \tag{A.2}$$

   (The second difference in the test statistic re-centers the bootstrapped CDF against the main estimate of the CDF.)

4. Repeat steps 1-3 many times using B replications ($b = 1...B$).

The main estimate for the difference between the two CDFs of interest is as follows:[1]

$$\bar{S}^{\text{Unequal, Home}} = \sup_z \left(F(z \,|\, \text{Home}) - F(z \,|\, \text{Unequal}) \right) \tag{A.3}$$

[1]Note that Barrett and Donald (2003) present the test statistic with an additional factor depending on the relevant sample sizes only. We omit this factor here because it applies symmetrically to the bootstrap samples and the main test statistic; this does not affect the result.

The resulting bootstrap p-value then is

$$\hat{p}^{\text{Unequal, Home}} = \frac{1}{B}\sum_{b=1}^{B} 1\left[\bar{S}_b^{\text{Unequal, Home}} > \bar{S}^{\text{Unequal, Home}}\right]. \tag{A.4}$$

For the selection-corrected estimates, we apply the same procedure, using the Heckman selection correction to compute the coefficient estimate $\hat{\beta}^{\star}$.[2]

A.3 Addressing Attrition

As described in the main text, the response rate to the follow-up survey is around 66 percent, overall, and about 59 percent for graduates who worked abroad at the time of the initial survey. To investigate whether differential attrition affects our findings, we re-estimate results for the full sample by carrying forward the responses from the initial survey and imputing some responses for individuals who did not respond in the follow-up survey. The imputation method follows three steps:

1. Carry forward responses from the initial survey:

   - Most variables that we use in the Mincer regression (the exact *university, university subject, university grades, studying abroad, ERASMUS/Total students in subject, pre-university mobility, school grades, apprenticeship, gender, parental occupation/education*) are collected in the initial survey. We therefore observe these variables even for individuals who do not reply to the follow-up survey.

2. Imputation of missing follow-up survey answers for individuals who reply in the initial survey but do not reply in the follow-up (or have missing values for the second round survey questions):

   - We impute the follow-up survey values for *marital/partner status, children, PhD/further degree completion* by using mean responses from individuals who respond in the follow-up survey, conditional on their initial survey response to the same question (i.e. we use the sample average corresponding to $x^{follow-up}_{non-respondents} = E\left(x^{follow-up}_{respondents}|x^{initial}_{response}\right)$). E.g. for marital status we impute the mean of being married during the follow-up

[2]The sampling procedure sometimes results in bootstrap samples where an institution is only represented by stayers. In the case of the selection-corrected estimates, these observations drop out of the sample in the probit stage because we include institution fixed effects. The resulting bootstrap samples therefore tend to be slightly smaller than the main sample in the case where we correct for sample selection.

survey, conditional on whether the individual was married in the initial survey or not.

- We impute the follow-up survey values for *potential experience* by using mean responses from individuals who respond in the follow-up survey (i.e. $x^{follow-up}_{non-respondents} = \bar{x}^{follow-up}_{respondents}$).
- We impute the follow-up survey values for the *country of work* by using the country of work information from the initial survey round (i.e. $x^{follow-up}_{non-respondents} = x^{initial}_{non-respondents}$), i.e. we assume that individuals remain in the country where they worked during the initial survey.[3]

3. Estimate the augmented Mincer regression for all individuals who work in Germany and respond to the follow-up survey.

4. Predict earnings for all individuals who respond to the initial survey.

Results are shown in Online Appendix Figure A.3 and in Online Appendix Table A.6.

A.4 Sensitivity of Results to Alternative Inequality Measures

In this section, we investigate the sensitivity of our main results to using alternative measures of inequality. The results are shown in Online Appendix Figure A.8. They focus on the three-group comparison with correction for sample selection in the Mincer regression. Panel (a) shows the main results, which are based on the 75th to 25th graduate earnings ratio. In Panels (b) to (f) we use different measures to classify countries. As an alternative measure of upper-tail inequality, we classify countries according to the *90/50 ratio* of the overall population. The results are essentially unchanged (panel (b)). In a second robustness check, we measure inequality with the 75/25 ratio of the *overall population*, which allows to measure inequality in larger samples. The resulting CDFs are very similar (panel (c)); if anything, the differences across the three groups are slightly more pronounced when we use this broader measure to classify countries. In a further test, we show that the results are similar if we use the 90/10 inequality ratio for the overall population as an alternative way of measuring inequality (panel (d)). We then show that the results also remain largely unchanged if we classify countries according to the Gini coefficient reported

[3]Because the initial survey of the 1993 cohort did not report the country of work (but only whether individuals worked abroad or not) we can only use individuals who worked in Germany during the initial survey round for the imputation of the 1993 cohort.

by the OECD, and thus using data from a completely different data source (panel (e)). Finally, we also obtain very similar results if we group countries based on the Theil index (panel (f)). We report stochastic dominance tests in Online Appendix Table A.8.

B Data Appendix

B.1 Construction of Inequality Measures

B.1.1 Data Sources on Earnings in Germany and Destination Countries

We collect data on 75/25 earnings differentials for the main destinations from the Luxembourg Income Study (LIS) (2013). Two important destinations for German university graduates, Austria and Switzerland, are not comprehensively covered in the LIS. We therefore use additional datasets for these countries: for Austria, the *Microcensus* (1999) and the *European Union Statistics on Income and Living Conditions (EU-SILC)* (2007 and 2008), and for Switzerland, the *Schweizerische Arbeitskräfteerhebung (SAKE)* (1998-2005) to collect data on additional years. Table A.3 summarizes the data sources, available survey years, and the number observations in each survey and year.

We then construct earnings inequality measures for the period 1998 to 2010.[4] We restrict the samples to full-time employed men and women between 30 and 60 years, exclude self-employed individuals, individuals who are still in school, and individuals who report zero or negative earnings. We apply the sampling weights of the surveys to calculate all statistics. Using these samples, we then construct earnings percentiles based on personal annual labor income.

To compare wage levels across countries, we convert each currency to U.S. dollars adjusted by purchasing-power-parity (ppp) measures from the Penn World Table (Heston et al., 2012). To express earnings in constant prices, we use the U.S. consumer price index for urban consumers from the U.S. Bureau of Labor Statistics (2013). These adjustments do not affect our inequality measures because inequality measures are based on percentile ratios and all adjustments cancel out when we compute ratios. Figure B.1 shows mean earnings by country and year for university graduates. Each dot represents one underlying survey in the LIS data and our additional data sources. As indicated by the figure, some surveys report gross earnings, while others report net earnings.

B.1.2 Constructing Measures of Net Earnings Using OECD Tax Data

We construct a consistent time series of net earnings by converting gross earnings into net earnings using tax rates from the OECD (2013c). The OECD reports three

[4]We calculate yearly earnings by multiplying monthly earnings by 12 for surveys that only report monthly earnings. The Austrian Microcensus only reports total income from all sources; we use this income measure to compute inequality.

tax rates for different positions in the earnings distribution. Tax rates for individuals with earnings equal to 67 percent, 100 percent, and 167 percent of average earnings.[5] From 2000 to 2010, the OECD reports tax rates for average workers (AW).[6] Before 2000, the OECD only reports tax rates for average production workers (APW).[7] As the definition of average workers includes white collar workers, average worker tax rates are closer to tax rates paid by university graduates. We construct tax rates for average workers (AW) for 1998 and 1999 using data from 2000 to 2004 - a period when the OECD reported tax rates for both APWs and AWs.[8] First, we regress the tax rate for AWs on the tax rate for APWs including country and time fixed effects for the period 2000 to 2004. We then use the estimated coefficients to predict tax rates for AWs for 1998 and 1999. Figure B.2 reports tax rates for workers with earnings equal to 67 percent, 100 percent, and 167 percent of average earnings.

The OECD takes into account that countries have different layers of taxes (see OECD, 2011, Part IV. Methodology and Limitations, pp. 561-566 for more details, and also OECD, 2001). In our sample nine countries (Australia, Austria, France, Germany, Ireland, Luxembourg, the Netherlands, Poland, and the UK) only have federal income taxes, three countries (Canada, Switzerland, and the United States) also have state income taxes, and nine countries (Belgium, Denmark, Finland, Italy, Japan, Norway, Sweden, Switzerland, and the United States) also have local income taxes. Spain has a different tax scheme for the Autonomous Regions. Depending on the country, taxes at the different layers are organized rather differently. In some countries (Belgium, Canada - other than Quebec, Denmark, Italy, Norway, and Spain), local taxes are a percentage of taxable income or tax paid to the central government. Other countries (Finland, Japan, Sweden, and Switzerland) offer different tax reliefs from central government taxes. Lastly, U.S. states have discretion over both the tax base and tax rates. The OECD considers sub-national tax rates

[5]We use tax rates for single persons without children because some surveys in the LIS data do not provide coherent information about household compositions. The majority of university graduates in the graduate cohort data are not married and do not have children. For the minority of graduates who are married and/or have children, tax rates for single persons without children may be too high. Nevertheless, these tax rates give a good indication of the general tax incidence in a country.

[6]The average worker (AW) is defined as "an adult full-time worker in the private sector whose wage earnings are equal to the average wage earnings of such workers. This definition includes manual and non-manual workers, supervisory workers as well as managerial workers" (OECD, 2013a).

[7]The average production worker (APW) is defined as "an adult full-time worker directly engaged in a production activity within the manufacturing sector whose earnings are equal to the average wage earnings of such workers. This definition includes manual workers and minor shop-floor supervisory workers. White collar workers are excluded." (OECD, 2013a).

[8]During this period, the two series are highly correlated (0.94 for the tax rate of 100 percent of average earnings).

in different ways: For some countries, the OECD assumes that the average worker (AW) lives in a typical area. This assumption is applied to Canada (AW lives in Ontario), Italy (AW lives in Rome), Switzerland (AW lives in canton and commune of Zurich) and the United States (AW lives in Detroit, Michigan). For other countries (Denmark, Finland, and Sweden) the OECD considers the cross-region average of sub-central government income taxes. For Japan, Spain, and Belgium the OECD considers the most commonly applied sub-national rates. Lastly, for Norway, the local rates do not vary within the country.

B.1.3 Construct 75/25 Differentials Based on Net Earnings

We construct earnings percentiles based on net earnings for each country and year between 1998 to 2010 by linearly interpolating percentiles for years with missing data. At endpoints, we extrapolate using the same value as in the last observed survey.[9] Table B.1 reports mean earnings, 25th, 50th, and 75th earnings percentiles for each country. We classify countries into either more or less equal destinations using average 75/25 differentials for the time period 1998 to 2010.

B.1.4 Construction of Theil Indices

For a robustness check, we categorize countries according to the Theil index (Figure A.8). Unlike Gini coefficients, Theil indices are not readily available from official sources for our time period and set of countries. We therefore compute Theil indices with data from Luxembourg Income Study (LIS) (2013) and EU-SILC data for Austria.

Schroeder and Boenke (2012) construct Theil indices with the 2000 wave of the LIS data for a selected sample of countries. As we need to measure inequality for the years 1998 to 2010 we apply the Schroeder and Boenke (2012) method to construct Theil indices for this longer time period and an extended set of countries. Following Schroeder and Boenke (2012), we restrict the sample in the income surveys to households that have at least one member who receives labor income and adjust disposable household income by dividing household income with the square root of the number of household members. We then drop the top and bottom 1 percent of households to exclude outliers. Finally, we construct the Theil index for the overall population in each country by multipling household weights by the number of household members and normalizing them to one for each country.

[9]For countries that report data for the pre-1998 period, we use the information in these early surveys to linearly interpolate between the last pre-1998 survey and the first post-1998 survey to obtain percentiles for years until the first available post-1998 survey.

B.1.5 Comparison of Average Wages

To verify the reliability of the augmented LIS data, we compare mean earnings (for the overall population) with official statistics from the OECD. In Figure B.3, we plot gross earnings averaged over the period from 1998 to 2010 against mean earnings from the OECD. The OECD data are ppp-adjusted and denoted in 2013 U.S. dollars and 2013 constant prices. Average annual earnings are computed per full-time equivalent dependent employee. The number is obtained by dividing the national-accounts-based total wage bill by the average number of employees in the total economy, which is then multiplied by the ratio of average usual weekly hours per full-time employee to average usually weekly hours for all employees. As is evident from the figure, the correlation between the two series is very high ($r = 0.899$).

B.2 Data on Confounding Factors

Table B.2 shows country-level values used in section 4.1. Unemployment rates are unemployment rates of 25-64 year-olds with tertiary education. Data come from the OECD and are taken from two editions of Education at a Glance. For the years 1998 to 2009, we take the 2011 edition (Table A7.4a, p. 131 and 132) and for the year 2010, we take the 2014 edition (Table A5.2a, p. 117). *Family expenditure* is public expenditure on family benefits, such as child allowances and credits, childcare support, income support during leave, and sole parent payments, as percent of GDP and it is taken from the OECD Social Expenditure Database (SOCX). We have yearly data on both series and take simple averages over the years 1998 to 2010. *Life satisfaction* is a component of the OECD *Better Life Index 2014* and is an average score which considers people's evaluation of their life as a whole. It is a weighted sum of diffwerent response categories based on people's rates of their current life relative to the best and worst possible lives for them on a scale from 0 to 10, using the Cantril Ladder (known also as the "Self-Anchoring Striving Scale"). The reference year is 2013 for all countries with the exception of 2012 for Norway, Switzerland, and the United States, and 2011 for Japan.

B.3 German Immigrants in the American Community Survey (ACS)

To identify high-skilled migrants from Germany in the United States, we use data from the American Community Survey (ACS). The data come from the Integrated Public Use Microdata Series (IPUMS) from the Minnesota Population Center (Rug-

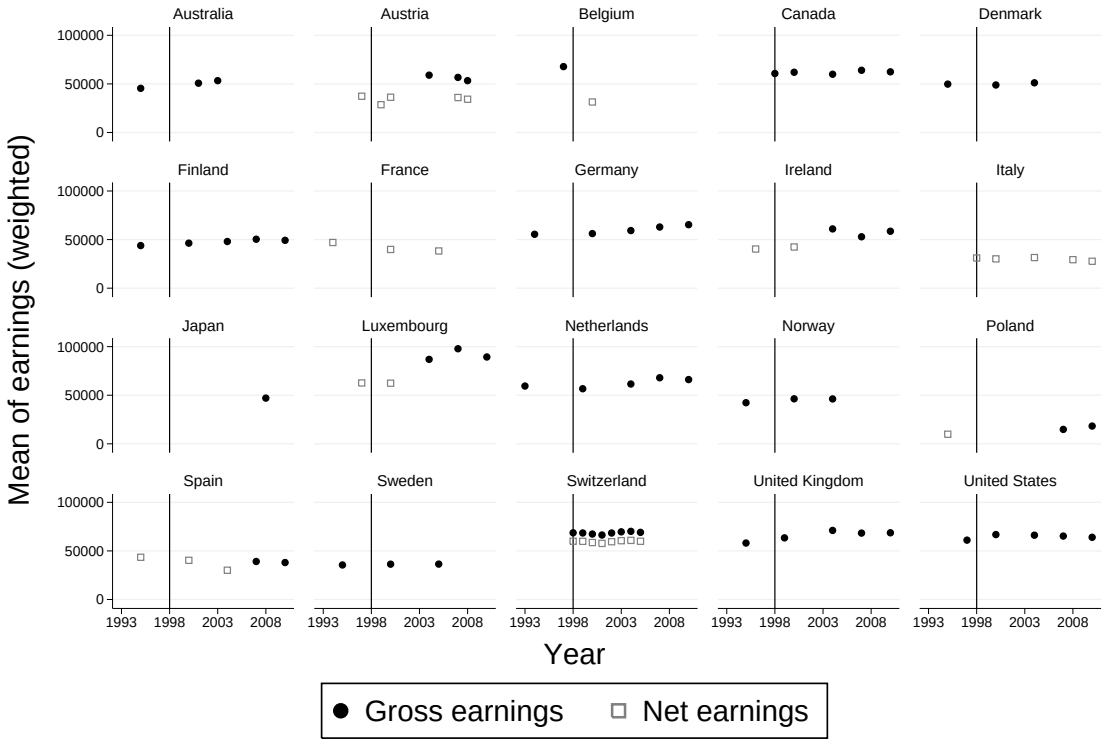
gles et al., 2010). The ACS is an administrative, yearly, and cross-sectional survey designed to collect representative information on the U.S. population between the decennial Censuses . It covers 1 percent of the U.S. population. To maximize sample size, we use the 2011 ACS three-year sample, which pools the 2009, 2010, and 2011 ACS waves. ACS waves before 2009 did not report the field of study, which is crucial for the estimation of the Mincer regression.

We restrict the sample to individuals who were either born in the United States or in Germany. To make German immigrants as similar as possible to the emigrants we study in our main results, we drop German immigrants whose parents are U.S. citizens and further restrict the sample to immigrants who migrated between 1996 and 2010 and were at least 25 years old when they migrated to the United States.

We further restrict the sample to individuals between 30 and 45 years of age, who are full-time employed (work between 50 and 52 weeks per year and more than 35 hours a week) and have at least a (four-year) bachelor's degree. These sample restrictions make the ACS sample as comparable as possible to the graduates who we observe in the German graduate survey. To identify individuals who are actually employed we further restrict the sample to individuals who report non-zero earnings. We keep earnings observations that are imputed (hot deck imputation) by the U.S. Census Bureau. After these restrictions, there are a few cases who report very low annual earnings and we therefore drop the lowest 1 percentile of the remaining earnings distribution. To mimic our analysis of emigrants from Germany, we assign German immigrants and U.S. natives to synthetic graduate cohorts: those between 42 and 45 years of age to graduate cohort 1993, those between 38 and 41 years of age to graduate cohort 1997, those between 34 and 37 years of age to graduate cohort 2001, and those between 30 and 33 years of age to graduate cohort 2005. The final estimation sample includes 289,538 U.S. natives (Column (1) of Table A.10) and 565 German immigrants (Column (2) of Table A.10).

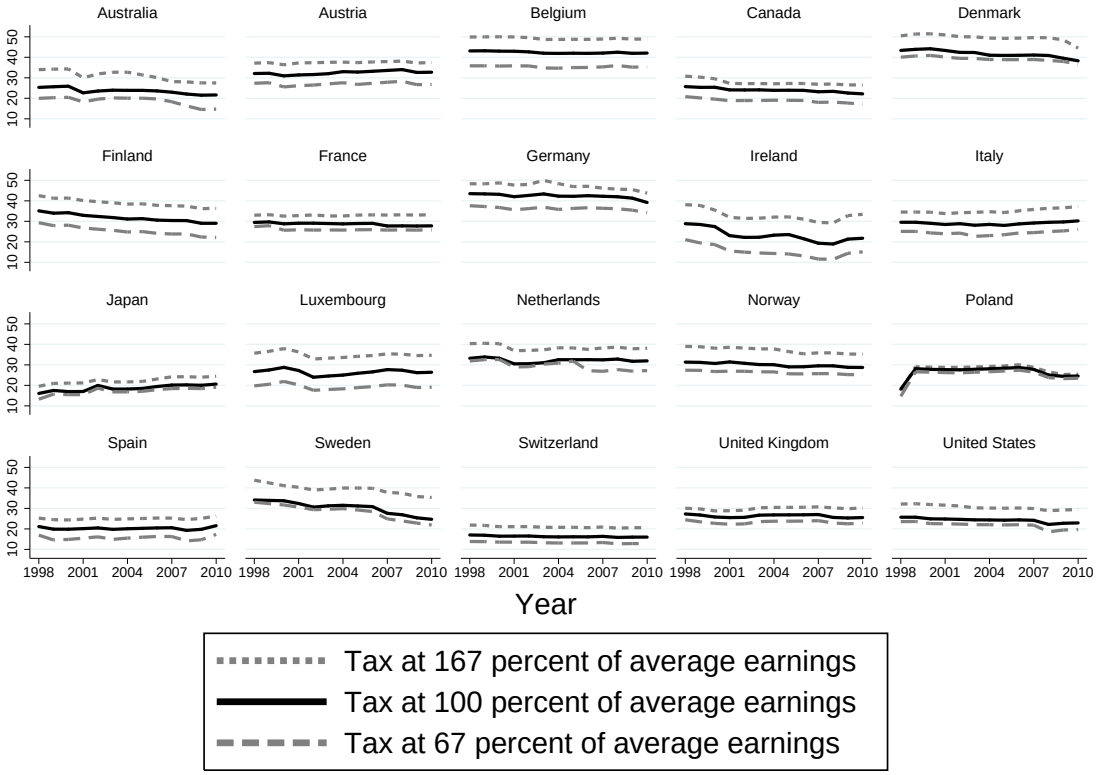
B.4 Figures and Tables

Figure B.1: Mean of weighted gross and net earnings in ppp-adjusted 2005 U.S. dollars



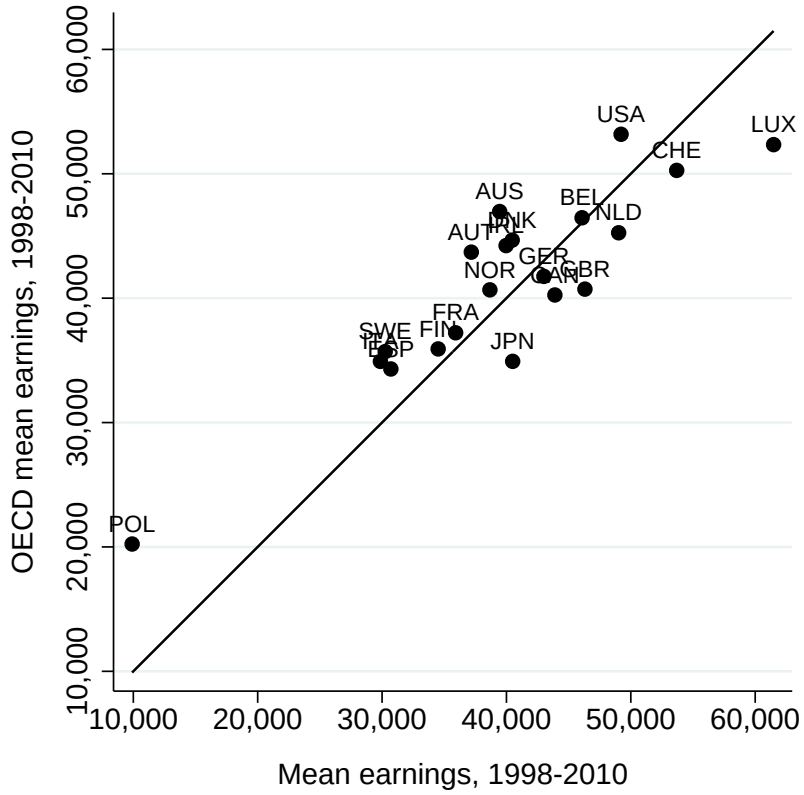
Notes: The figure shows mean values of weighted gross and net earnings in ppp-adjusted 2005 U.S. dollars over countries and years for university graduates. We use the U.S. consumer price index to convert current dollars into constant 2005 dollars (U.S. Bureau of Labor Statistics, 2013). Purchasing-power-parity measures are from Heston et al. (2012). Pre-2002 earnings are adjusted to Euro values for Euro countries. For each country, the sample is restricted to full-time employed university graduates between 30 and 60 years. We only consider regular or dependent employment. Data are collected from the Luxembourg Income Study (LIS) (2013) and from additional income surverys for Austria and Switzerland. The horizontal line indicates the year 1998, the first year for which we observe German university graduates.

Figure B.2: Tax rate series



Notes: The figure shows the time series for the tax data in each country. Data come from OECD (2013c).

Figure B.3: Comparison of mean earnings



Notes: The figure plots our computed measure for average mean gross earnings between 1998 and 2010 against average earnings obtained from the OECD. The black line is a 45° line.

Table B.1: Adjusted earnings percentiles

		Uni graduates		All	
		net	gross	net	gross
Australia	Mean	39,751	52,066	30,122	39,456
	P25	28,038	34,893	20,598	25,635
	P50	34,485	45,171	25,974	34,023
	P75	40,368	59,344	30,907	45,429
Austria	Mean	35,804	54,016	25,236	54,016
	P25	25,117	35,005	17,476	35,005
	P50	32,197	48,124	22,789	48,124
	P75	41,195	65,211	29,879	65,211
Belgium	Mean	31,891	55,952	26,264	46,081
	P25	22,755	35,411	19,874	30,929
	P50	27,567	48,366	23,745	41,662
	P75	34,966	69,947	29,145	58,307
Canada	Mean	47,078	62,085	33,225	43,804
	P25	28,832	35,599	20,026	24,719
	P50	39,516	52,122	27,823	36,689
	P75	50,309	69,814	39,132	54,309
Denmark	Mean	29,181	50,405	23,433	40,464
	P25	21,640	35,865	17,764	29,435
	P50	25,475	44,000	21,567	37,243
	P75	29,140	58,445	23,356	46,826
Finland	Mean	32,061	47,343	22,561	33,300
	P25	23,966	32,415	17,394	23,509
	P50	28,461	42,025	19,895	29,369
	P75	33,477	55,351	23,207	38,354
France	Mean	39,151	55,081	25,529	35,915
	P25	24,810	33,471	17,097	23,064
	P50	34,007	47,845	21,868	30,764
	P75	46,823	69,796	29,673	44,225
Germany	Mean	34,791	59,957	24,947	43,010
	P25	25,160	39,269	18,322	28,606
	P50	30,932	53,316	22,482	38,771
	P75	38,368	72,562	27,069	51,212

(continued on next page)

Table B.1 (continued)

		Uni graduates		All	
		net	gross	net	gross
Ireland	Mean	44,052	57,869	30,439	39,969
	P25	31,521	37,358	20,999	24,886
	P50	39,708	52,270	26,982	35,452
	P75	47,785	71,804	33,430	50,193
Italy	Mean	30,155	42,589	21,141	29,866
	P25	19,849	26,270	16,219	21,472
	P50	25,230	35,630	19,211	27,143
	P75	35,840	55,376	23,434	36,219
Japan	Mean	37,556	47,087	32,303	40,500
	P25	25,839	31,719	19,379	23,790
	P50	35,293	44,249	30,359	38,063
	P75	45,190	59,633	40,382	53,289
Luxembourg	Mean	65,621	90,028	44,816	61,482
	P25	46,929	58,708	27,066	33,858
	P50	59,277	81,329	39,676	54,446
	P75	72,803	113,245	51,769	80,598
Netherlands	Mean	40,306	60,130	31,733	47,347
	P25	27,980	40,814	22,948	33,476
	P50	34,487	51,470	27,529	41,085
	P75	41,390	67,909	32,284	52,966
Norway	Mean	32,095	46,065	26,943	38,670
	P25	23,621	32,217	19,446	26,522
	P50	28,590	41,031	24,827	35,631
	P75	33,266	53,576	28,410	45,755
Poland	Mean	9,886	12,305	6,660	8,290
	P25	6,090	7,400	4,325	5,255
	P50	8,283	10,310	5,653	7,036
	P75	11,628	14,619	7,777	9,777
Spain	Mean	33,985	42,671	24,453	30,708
	P25	22,141	26,295	16,173	19,207
	P50	30,444	38,231	21,049	26,435
	P75	39,221	52,311	28,978	38,659

(continued on next page)

Table B.1 (continued)

		Uni graduates		All	
		net	gross	net	gross
Sweden	Mean	24,690	36,377	20,540	30,251
	P25	17,143	24,519	14,747	21,086
	P50	21,627	31,851	18,819	27,710
	P75	25,100	42,143	21,002	35,254
Switzerland	Mean	59,760	68,756	46,610	53,688
	P25	44,463	50,950	33,867	38,996
	P50	54,739	62,943	42,039	48,333
	P75	68,951	79,418	54,113	62,264
UK	Mean	49,582	67,625	33,958	46,309
	P25	30,733	40,239	20,011	26,199
	P50	39,983	54,535	27,569	37,600
	P75	53,003	76,090	37,356	53,629
United States	Mean	49,502	65,442	37,234	49,217
	P25	27,432	35,125	19,474	24,935
	P50	38,635	51,074	28,976	38,303
	P75	52,929	76,303	40,306	58,097

Notes: The table shows adjusted earnings percentiles by country. The data are restricted to 1998–2010 and include inter- and extrapolation between the years. The measures are denoted in ppp-adjusted 2005 U.S. dollars.

Table B.2: Country-level control variables

Country		(1)	(2)	(3)
Name	Label	Family expenditure	Unemployment rate	Life satisfaction
Australia	AU	2.82	2.89	7.4
Austria	AT	2.78	2.08	7.5
Belgium	BE	2.66	3.42	7.1
Canada	CA	1.08	4.65	7.6
Denmark	DK	3.72	3.57	7.6
Finland	FI	3.03	4.31	7.4
France	FR	3.03	5.23	6.7
Germany	DE	2.08	4.45	7.0
Ireland	IE	2.85	2.90	6.8
Italy	IT	1.35	5.43	6.0
Japan	JP	0.79	3.29	6.0
Luxembourg	LU	3.54	3.06	7.1
Netherlands	NL	1.64	2.12	7.4
Norway	NO	3.05	1.77	7.7
Poland	PL	1.13	4.62	5.7
Spain	ES	1.13	8.06	6.2
Sweden	SE	3.34	3.78	7.4
Switzerland	CH	1.22	2.27	7.8
United Kingdom	GB	3.12	2.46	6.9
United States	US	0.75	2.89	7.0

Notes: Family expenditure is public expenditure on family benefits as a percent of GDP and *unemployment rate* is the unemployment rates of 25-64 year-olds with tertiary education. Both series are simple averages for the years 1998 to 2010. Life satisfaction is an average score and considers people's evaluation of their life as a whole. The reference year is 2013 for all countries with the exception of 2012 for Norway, Switzerland, and the United States and 2011 for Japan. Data come from the OECD. See section **??** for details.

Appendix References

Barrett, G. F. and Donald, S. G. (2003). Consistent Tests for Stochastic Dominance. *Econometrica*, 71(1):71–104.

Chernozhukov, V., Fernández-Val, I., and Galichon, A. (2010). Quantile and Probability Curves without Crossing. *Econometrica*, 78(3):1093–1125.

Heston, A., Summers, R., and Aten, B. (2012). Penn World Table Version 7.1. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Luxembourg Income Study (LIS) (2013). Database (multiple countries). http://www.lisdatacenter.org.

OECD (2001). Taxing Wages 2000. OECD Publishing. http://dx.doi.org/10.1787/tax_wages-2000-en-fr.

OECD (2011). Taxing Wages 2010. OECD Publishing. http://dx.doi.org/10.1787/tax_wages-2010-en.

OECD (2013a). Glossary of Statistical Terms. http://stats.oecd.org/glossary/.

OECD (2013c). OECD Tax Statistics, Taxing Wages. http://dx.doi.org/10.1787/ctpa-twg-data-en.

Ruggles, S., Alexander, J. T., Genadek, K., Goeken, R., Schroeder, M. B., and Sobeku, M. (2010). Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database].

Schroeder, C. and Boenke, T. (2012). Country Inequality Rankings and Conversion Schemes. *Economics: The Open-Access, Open-Assessment E-Journal*, 6:2012–2028.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.

U.S. Bureau of Labor Statistics (2013). Consumer Price Index (CPI) - All Urban Consumers. http://www.bls.gov/cpi/data.htm.